L'Actualité économique

Une condition d’invariance du modèle de régression à coefficients aléatoires

Gordon Fisher

Volume 80, numéro 2-3, juin–septembre 2004

Hommage à Marcel Dagenais

URI : https://id.erudit.org/iderudit/011393ar
DOI : https://doi.org/10.7202/011393ar

Éditeur(s)

HEC Montréal

ISSN

0001-771X (imprimé)
1710-3991 (numérique)

Résumé de l'article

Cet article développe une condition d’invariance du modèle de régression à coefficients aléatoires (RCA) et de cette façon permet : (i) d’étendre les résultats de Rao (1963, 1967)”; (ii) de fournir une nouvelle démonstration de l’égalité des estimateurs des moindres carrés généralisés directs et ceux en deux étapes du modèle RCA”; et (iii) de rectifier une proposition de McAleer (1992) pour les modèles mixtes.

L'Actualité économique, Revue d'analyse économique, vol. 80, nos 2-3, juin-septembre 2004

UNE CONDITION D'INVARIANCE DU MODÈLE DE RÉGRESSION À COEFFICIENTS ALÉATOIRES*

Gordon FISHER
Université Concordia

RÉSUMÉ – Cet article développe une condition d'invariance du modèle de régression à coefficients aléatoires (RCA) et de cette façon permet : (i) d'étendre les résultats de Rao (1963, 1967); (ii) de fournir une nouvelle démonstration de l'égalité des estimateurs des moindres carrés généralisés directs et ceux en deux étapes du modèle RCA; et (iii) de rectifier une proposition de McAleer (1992) pour les modèles mixtes.

ABSTRACT – This paper develops an invariance condition for the random coefficients regression (RCR) model and thereby: (i) exends the results originally proposed by Rao (1965, 1967); (ii) provides a new proof of the equality of the direct and two-step estimators of the RCR model; and (iii) corrects a result claimed in McAleer (1992) for mixed models.

INTRODUCTION

Au début des années quatre-vingt-dix, Marcel Dagenais, en coopération avec Jean-Marie Dufour, a publié quatre articles portant sur l'invariance ou, pour être plus précis, sur l'absence de l'invariance dans certains tests de modèles non linéaires lors d'un changement des unités de mesure. Il semble donc à propos d'honorer la contribution de Marcel en économétrie et sa carrière en tant qu'économètre en produisant un article au sujet de l'invariance. Toutefois, on ne traitera pas de l'absence d'invariance de certains tests statistiques dans différents contextes mais plutôt de l'invariance d'un sous-espace de régression par rapport à la dispersion de ses erreurs. Dans ces circonstances, les estimateurs des moindres carrés ordinaires (MCO) des coefficients fixes d'une régression linéaire sont identiques aux estimateurs des moindres carrés généralisés (MCG) correspondants. Ainsi, les conditions admettant ce type d'invariance permettent de caractériser la structure de la dispersion des erreurs « qui peut être ignorée sans conséquence à la meilleure estimation linéaire sans biais » (Baksalary, 1988 : 98).

* L'auteur voudrait remercier Sana Mohsni pour son aide durant la préparation de cet article. Il reconnaît aussi l'appui financier des bourses no SOO462 du CRSHC et no RO-20-B245 de l'Université Concordia.

Dans le présent article, l'invariance sera étudiée dans le cadre du modèle de régression à coefficients aléatoires (RCA) et ce, par opposition au modèle de régression à coefficients fixes (RCF). Ce dernier a été largement employé dans la littérature; par exemple, Puntanen et Styan (1989), en prétendant fournir « une liste complète de références », citent non moins de 86 travaux qui ont suivi la première contribution d'Anderson (1948). Cette liste n'inclut pas les articles parus à partir de 1990 ainsi que plusieurs papiers en économétrie parus avant et après 1989; dont : Balestra (1970), Philoche (1971), Dwivedi et Srivastava (1978), Gouriéroux et Monfort (1980), Kapteyn et Fiebig (1981), Kloek (1981), Srivastava et Giles (1987), Baksalary et Trenkler (1989), et McAleer (1992).

En effet, la contribution de Puntanen et Styan (1989) était adressée aux statisticiens et ainsi n'a pas inclus des applications économétriques. Or, les problèmes d'invariance en économétrie sont plus intéressants que ceux de la statistique classique. Ceci est dû à l'omniprésence de l'endogénéité dans les modèles économétriques, à la nécessité concomitante de formuler des systèmes d'équations interdépendants et au besoin de permettre la variation des coefficients dans le temps ou pour différentes unités ou politiques économiques (voir Swamy *et al.* 1988a, 1988b et 1989). Cet article s'intéresse aux variations des coefficients dans le cadre des modèles RCA. Néanmoins, on peut décrire brièvement le problème principal dans le contexte du modèle RCF.

Supposons qu'il existe un modèle RCF standard avec une dispersion des erreurs Ω définie positive. Ω peut être exprimée comme la somme de deux dispersions séparées Σ et Λ. Σ est une matrice définie positive et Λ peut être définie non négative; chaque matrice représentant une source, séparée et linéairement indépendante, de variation des erreurs. La question à poser est formulée comme suit : sous quelles conditions l'estimateur MCG des coefficients de la régression linéaire sera-t-il indépendant de Λ? En recherchant de telles conditions, Σ n'est pas contrainte à être un multiple positif de la matrice identité comme dans Rao (1965, 1967). Rao a été le premier à développer les conditions selon lesquelles l'estimateur MCG pour ce problème est indépendant de Λ. Ainsi, les résultats de cet article sont plus généraux que ceux disponibles. Ces résultats sont particulièrement intéressants car ils peuvent être résumés sous une nouvelle condition d'invariance qui est simple à appliquer et facilement repérable.

L'article s'inspire de deux volets de la littérature : l'invariance dans un contexte RCF, déjà mentionnée, et la contribution au modèle RCA. Swamy (1971) présente les contributions au modèle RCA jusqu'au début des années soixante-dix. Des revues de cette littérature plus à jour sont disponibles dans Pfefferman (1984), Swamy *et al.* (1988a, 1988b, 1989) et Longford (1993). Une contribution notable à cette littérature, et d'une importance particulière ici, a été faite par Rao (1965).

Le volet de la littérature relié à l'invariance, qui trouve ses origines dans Anderson (1948), est souvent considéré comme ayant été initié et résolu par Zyskind (1962, 1967) et Rao (1967) avec des contributions significatives de Kruskal (1968), Watson (1967, 1972), Piloche (1971) et Gouriéroux et Monfort

(1980). Une grande partie de ces travaux ainsi que plusieurs applications économétriques ont été revues dans McAleer (1992). Puntanen et Styan (1989) ont fourni une revue approfondie de la littérature statistique.

Avant de faire état du contenu de l'article, qui fait partie des mémoires de Marcel, il est approprié d'ajouter un commentaire personnel. Une grande partie de l'article est expliquée en utilisant le langage et la méthode propres à l'algèbre sans coordonnées. Pendant les 20 ans où j'ai connu Marcel, il a préféré considérer les mathématiques comme une science calculable, ce qui lui a permis d'éviter l'abstraction caractérisant l'approche sans coordonnées. Je dirais même que les arguments ayant recours à cette approche l'irritaient un peu. Or, on nous dit que le paradis ouvre les portes aux pensées les plus vastes et les plus sublimes. Je suis dès lors confiant que l'irritation terrestre, qu'il a pu ressentir, sera transformée en un sourire paradisiaque à la pensée que son vieil ami n'a pas abandonné ses vieux trucs.

Dans la prochaine section, la condition de l'invariance est établie par rapport à un modèle RCF ayant deux dispersions additives. Ceci est suivi dans la deuxième section de trois applications : la première traite d'un problème simple de régression linéaire avec des erreurs équicorrélées, la deuxième traite d'une estimation MCG d'un modèle RCA général tel que décrit dans l'article remarquable de Pfefferman (1984) et la dernière concerne un modèle linéaire mixte avec à la fois des coefficients fixes et des coefficients stochastiques, tel que décrit par McAleer (1992).

Un problème connexe qui est étudié dans cet article est un sous-produit logique de la condition de l'invariance établie dans la première section. Ceci revient à fournir une explication géométrique de l'équivalence entre la méthode MCG directe et celle en deux étapes dans l'estimation de la moyenne des coefficients d'un modèle RCA général. Cette équivalence est souvent montrée par une manipulation de matrices, facile à effectuer mais qui n'offre aucune intuition et donc n'aide pas à la compréhension des résultats générés. Une explication géométrique de l'équivalence entre la méthode MCG directe et celle en deux étapes est développée dans la troisième section. Les résultats sont résumés dans la dernière section.

1. L'INVARIANCE ET LES DISPERSIONS ADDITIVES

La notation $q \sim [\alpha, B]$ est utilisée pour définir un vecteur aléatoire q distribué avec une moyenne $\mathrm{E}[q] = \alpha$ et une dispersion $\mathrm{E}\left[(q - \alpha)\,(q - \alpha)^{\top}\right] = B$. $\mathbf{Q}^m$ représente l'ensemble de matrices définies positives dans $\mathbf{R}^m$ et $\mathbf{N}^m$ l'ensemble de matrices définies non négatives dans $\mathbf{R}^m$; bien entendu, $\mathbf{Q}^m \subset \mathbf{N}^m$.

Le cadre à étudier comprend un vecteur aléatoire y qui varie dans $\mathbf{R}^n$ selon $y \sim [X\,\beta, (\Sigma + \Lambda)]$, X étant une matrice fixe et connue de dimension $(n \times k)$ et de rang k et β un vecteur de dimension $(k \times 1)$ et de coefficients inconnus; $\Sigma \in \mathbf{Q}^n$, $\Lambda \in \mathbf{N}^n$. En écrivant $\Sigma + \Lambda = \Omega$, alors

$$y = X\,\beta + \varepsilon, \quad \varepsilon \sim [0, \Omega], \; \Omega \in \mathbf{Q}^n. \tag{1}$$

Soit $X\beta = \mu$. Alors $\mu \in \mathcal{X}$, le sous-espace de régression de dimension $\dim \mathcal{X} = k$ défini par

$$\mathcal{X} = \{x \in \mathbf{R}^n : x = X\lambda, \lambda \in \mathbf{R}^k\}. \tag{2}$$

Dans ce contexte simple, trois estimateurs seront discutés. Premièrement, l'estimateur MCO μ^* de μ, $\mu^* = X\beta^*$, où

$$\mu^* = X(X^\top X)^{-1} X^\top y = Py. \tag{3}$$

P étant la matrice de projection sur $\mathcal{X}$ orthogonale au produit scalaire $(\cdot, \cdot)$; ainsi, $P = X(X^\top X)^{-1} X^\top$ et $\mu^* \in \mathcal{X}$ est la solution unique à

$$(y - \mu^*, x) = (y - \mu^*)^\top x = 0 \quad \forall x \in \mathcal{X}. \tag{4}$$

Le deuxième estimateur est l'estimateur MCG unique $\widehat{\mu} \in \mathcal{X}$ de μ défini par

$$(y - \widehat{\mu}, \Omega^{-1} x) = (y - \widehat{\mu})^\top \Omega^{-1} x = 0 \quad \forall x \in \mathcal{X}. \tag{5}$$

La solution à (5) est $\widehat{\mu} = X\widehat{\beta} = P_\Omega y = X(X^\top \Omega^{-1} X)^{-1} X^\top \Omega^{-1} y$, P_Ω étant la matrice de projection sur $\mathcal{X}$ orthogonale au produit scalaire $(\cdot, \Omega^{-1} \cdot)$. Le troisième estimateur est un estimateur des moindres carrés intermediaires (MCI) $\widetilde{\mu} \in \mathcal{X}$ de μ, défini par

$$(y - \widetilde{\mu}, \Sigma^{-1} x) = (y - \widetilde{\mu})^\top \Sigma^{-1} x = 0 \quad \forall x \in \mathcal{X}. \tag{6}$$

L'estimateur MCI est $\widetilde{\mu} = X\widetilde{\beta} = P_\Sigma y = X(X^\top \Sigma^{-1} X)^{-1} X^\top \Sigma^{-1} y$, P_Σ étant la matrice de projection sur $\mathcal{X}$ orthogonale au produit scalaire $(., \Sigma^{-1}.)$. Notons que P, P_Ω et P_Σ sont toutes membres de la famille $\mathbf{P}$ de matrices de projections sur $\mathcal{X}$ et, alors, sont toutes idempotentes et appartenant à la famille $\mathbf{N}^n$ tel que, pour tout $F \in \mathbf{P}$, $Fx = x$ pour tout x en $\mathcal{X}$.

Le problème étudié originellement par Anderson (1948) et de manière plus approfondie par Zyskind (1962, 1967) et Rao (1967) s'intéresse aux conditions nécessaires et suffisantes pour établir l'égalité entre β^* et $\widehat{\beta}$ (ou bien μ^* et $\widehat{\mu}$). Kruskal (1968) a été le premier à formuler le problème dans un cadre libre de coordonnées et ses résultats ont été généralisés par Philoche (1971). Assez curieusement, le problème de l'égalité entre estimateurs a été considéré comme un problème statistique, mais la solution s'avère être entièrement algébrique. Le seul point d'intérêt statistique est la qualité de la dispersion des erreurs qui va générer l'égalité recherchée. La seule condition de base algébrique exploitée par Kruskal (1968), 26 ans aprés son apparition dans le livre de Halmos (1942), est le concept d'invariance tel que présenté ci-après :

Définition : *Soit la matrice A de dimension $(n \times n)$. Un sous-espace $\mathcal{M}$ de $\mathbf{R}^n$ est dit invariant sous A si, pour tout $x \in \mathcal{M}$, $Ax \in \mathcal{M}$.*

Étant donné le sous-espace $\mathcal{M}$ avec une matrice base M, soit $\mathcal{N}$ un sous-espace complémentaire de $\mathcal{M}$ dans $\mathbf{R}^n$ avec une matrice base N. Ainsi $\mathcal{M} \oplus \mathcal{N}$ et $[M : N] = \mathrm{R}$ qui est une base de $\mathbf{R}^n$. Puisque A transforme des vecteurs de $\mathbf{R}^n$ à $\mathbf{R}^n$, une implication immédiate de la matrice A qui est dépendante des coordonnées est

$$AR = RB \tag{7}$$

pour une matrice carrée B. En divisant B en quatre sous-matrices $B_{ij}(i, j = 1, 2)$ conformément à la répartition de R en M et N, il est immédiatement possible de montrer qu'une condition nécessaire et suffisante pour que $\mathcal{M}$ soit invariante sous A est

$$B_{21} = 0 \Leftrightarrow AM = MB_{11} \tag{8}$$

Halmos (1958 : 71-72). Si maintenant $\mathcal{N} = \mathcal{M}^{\perp}$, alors $M^{\top} N = 0$ et la solution de A en (7) sujet à (8) est $A = RBR^{-1}$ ou

$$A = MCM^{\top} + MDN^{\top} + NEN^{\top} + I_n \theta \tag{9}$$

pour des valeurs aléatoires de θ, les matrices C, D et E dépendant de θ, M, N, leurs inverses Moore-Penrose et la matrice restreinte B. En outre, étant donné (9), (7) peut être récupérée sujet à (8). Il s'ensuit qu'une condition nécessaire et suffisante pour que $\mathcal{M}$ soit invariante sous A est d'exprimer A sous la forme (9). Ainsi chacune des équations (8) et (9) représente des conditions nécessaires et suffisantes de l'invariance de $\mathcal{M}$ sous A qui dépendent des coordonnées.

Il est important de noter que, dans (8) et (9), A n'est pas tenue d'être symétrique, ce qui aurait pu être le cas si A était une matrice de dispersion des erreurs. La symétrie permet plus d'épuration. En particulier, lorsque $A = A^{\top}$, B de (7) devient une matrice diagonale par blocs et $B_{21} = 0$, $B_{12} = 0$; en plus, $\mathcal{M}$ et $\mathcal{M}^{\perp}$ sont toutes les deux invariantes sous A. Ainsi (8) devient

$$AM = MB_{11}, AN = NB_{22} \tag{10}$$

alors que, en (9), $MDN^{\top} = 0$ et

$$A = A^{\top} = MCM^{\top} + NEN^{\top} + I_n\theta. \tag{11}$$

Si, de plus $A \in \mathbf{N}^n$, alors C, E et θ doivent satisfaire cette contrainte additionnelle.

Si on revient au modèle de régression (1) et (2), le théorème principal de Kruskal (1968) peut être énoncé comme suit,

Théorème 1 : *Dans une régression linéaire définie par les équations* (1) *et* (2), *l'estimateur MCO* μ^* *de* (3) *et* (4) *est identique à l'estimateur MCG* $\widehat{\mu}$ *de* (5) *si et seulement si le sous-espace de régression* $\mathcal{X}$ *de* (2) *est invariant sous* Ω *de* (1) .

Le théorème 1 peut être mis en pratique en utilisant :

Théorème 2 : *Si un sous-espace* $\mathcal{M}$ *de* $\mathbf{R}^n$ *est invariant sous la matrice A de dimension* $(n \times n)$ *alors* $EAE = AE$ *pour toute projection E sur* $\mathcal{M}$. *Inversement, si* $EAE = AE$ *pour une projection E sur* $\mathcal{M}$, *alors* $\mathcal{M}$ *est invariant sous A.*

Démonstration : Halmos (1958 : 76) ■

Les théorèmes 1 et 2 montrent que $\mu^* = \widehat{\mu}$ si et seulement si $F\,\Omega\,F = \Omega\,F$ pour tout $F \in \mathbf{P}$; en particulier $P\,\Omega\,P = \Omega\,P \Leftrightarrow \Omega\,P = P\,\Omega$, par symétrie. Cette dernière condition est connue sous la *Condition de Zyskind* par Puntanen et Styan (1989). Bien sûr, $P\,\Omega\,P = \Omega\,P \Leftrightarrow Q\,\Omega\,P = 0$ où $Q = (I - P)$, la projection orthogonale dans $\mathcal{X}^{\perp} = \{z \in \mathbf{R}^n : (x, z) = 0\ \forall x \in \mathcal{X}\}$ et $Q\,\Omega\,P = 0 \Leftrightarrow P\,\Omega\,Q = 0 \Leftrightarrow \mathcal{X}^{\perp}$ est invariante sous Ω, par le théorème 2. De plus, $Q\,\Omega\,P = 0 \Leftrightarrow Z^{\top}\,\Omega\,X = 0$, où Z est une matrice base de dimension $n \times (n - k)$ de $\mathcal{X}^{\perp}$ ce qui implique que $Z^{\top}\,X = 0$. Finalement, allant de pair avec Rao (1967) et Zyskind (1967), $\mathcal{X}$ est invariant sous Ω si et seulement si, en utilisant l'équation (10), il existe une matrice carrée G tel que

$$\Omega\,X = XG \tag{12}$$

ou de la même manière, en utilisant (11), Ω doit satisfaire,

$$\Omega = XCX^{\top} + ZEZ^{\top} + I_n\,\theta \tag{13}$$

où θ, C et D sont arbitraires mais choisies pour satisfaire $\Omega \in \mathbf{Q}^n$.

En résumé, le théorème 2 et les équations (12) et (13) présentent des conditions nécessaires et suffisantes équivalentes pour établir l'égalité entre μ^* et $\widehat{\mu}$. Le théorème 1 représente une condition générale libre de coordonnées qui peut être implantée en appliquant le théorème 2 (qui est aussi libre de coordonnées) tandis que les équations (12) et (13) sont des conditions spécifiques dépendantes des coordonnées.

Les théorèmes 1 et 2 préparent le terrain pour déterminer les conditions sous lesquelles $\widehat{\mu}$ en (5), qui est l'estimateur MCG de μ, est identique à $\tilde{\mu}$ en (6) qui est l'estimateur ILS de μ.

Théorème 3 : *Dans une régression linéaire telle que définie en* (1)*, l'estimateur MCG* $\widehat{\mu}$ *de* (5) *est le même que l'estimateur ILS* $\tilde{\mu}$ *de* (6) *si et seulement si le sous-espace de régression* $\mathcal{X}$ *est invariant sous* $\Lambda\,\Sigma^{-1}$.

Démonstration : Soit $\widehat{\mu} = \tilde{\mu} = m$. Alors

$$(y - m, \Omega^{-1}\,x) = (y - m, \Sigma^{-1}\,x) = 0 \quad \forall x \in \mathcal{X}.$$

Maintenant $\Omega^{-1} = (\Sigma + \Lambda)^{-1} = (I_n + \Sigma^{-1}\,\Lambda)^{-1}\,\Sigma^{-1}$ d'où

$$(y - m, [I_n + \Sigma^{-1}\,\Lambda]^{-1}\,\Sigma^{-1}\,x) = (y - m, \Sigma^{-1}\,x) = 0 \quad \forall x \in \mathcal{X}.$$

Donc, $\Sigma^{-1}\,\mathcal{X}$ est invariant sous $(I_n + \Sigma^{-1}\,\Lambda)^{-1}$ et par conséquent sous $(I_n + \Sigma^{-1}\,\Lambda)$. Alors, $(I_n + \Sigma^{-1}\,\Lambda)\,\Sigma^{-1}\,x \in \Sigma^{-1}\,\mathcal{X}$ pour tout x en $\mathcal{X}$. Puisque $\Sigma \in \mathbf{Q}^n$, cette dernière expression implique que

$$x + \Lambda\,\Sigma^{-1}\,x \in \mathcal{X}, \quad \forall x \in \mathcal{X},$$

et donc que $\Lambda\,\Sigma^{-1}\,x \in \mathcal{X}$ pour tout $x \in \mathcal{X}$; ainsi $\mathcal{X}$ est invariant sous $\Lambda\,\Sigma^{-1}$.

Inversement, soit $\mathcal{X}$ invariant sous $\Lambda\,\Sigma^{-1}$. Donc pour tout $x \in \mathcal{X}$

$x + \Lambda\,\Sigma^{-1}\,x \in \mathcal{X}$

$\Rightarrow (\Sigma + \Lambda)\,\Sigma^{-1}\,x \in \mathcal{X}$

$\Rightarrow \Sigma^{-1}\,x \in \Omega^{-1}\,\mathcal{X}$

d'où il existe un m dans $\mathcal{X}$ tel que

$$(y - m, \Sigma^{-1}\,x) = (y - m, \Omega^{-1}\,x) = 0 \quad \forall x \in \mathcal{X}.$$

À gauche de l'égalité, le produit scalaire $m = \widehat{\mu}$ et à droite de l'égalité le produit scalaire $m = \widetilde{\mu}$. D'où, par l'unicité, $m = \widehat{\mu} = \widetilde{\mu}$. ■

Comme un corollaire au théorème 3, soit $\mathcal{X}$ invariant sous $\Lambda\,\Sigma^{-1}$ (mais pas invariant sous Σ). Alors, puisque $\Lambda\,\Sigma^{-1}$ n'est pas symétrique, une équation correspondante à (11) doit être développée. Ceci peut être fait en définissant W comme une matrice base dans le sous-espace $\mathcal{W} = \{w \in \mathbf{R}^n : (w, \Sigma^{-1}\,x) = 0\ \forall x\ \varepsilon\ \mathcal{X}\}$ qui est complémentaire, mais oblique à $\mathcal{X}$. Donc, une nouvelle base de $\mathbf{R}^n$ est $\mathrm{R} = [X : W]$ et $X^\top\,\Sigma^{-1}\,W = 0$. Maintenant, puisque $\mathcal{X}$ est invariant sous $\Lambda\,\Sigma^{-1}$ qui est $\mathbf{R}^n \mapsto \mathbf{R}^n$, $\Lambda\,\Sigma^{-1}\,\mathrm{R} = \mathrm{R}B$ pour une matrice carrée B qui peut être répartie comme avant en $B_{ij}(i, j = 1, 2)$ conformément à la répartition de R en X et W. Donc $\Lambda\,\Sigma^{-1} = \mathrm{R}B\mathrm{R}^{-1}$ et pour refléter l'invariance de $\mathcal{X}$ sous $\Lambda\,\Sigma^{-1}$, $B_{21} = 0$. Ainsi,

$$\begin{aligned}
\Lambda\,\Sigma^{-1} &= [X : W]\begin{bmatrix} B_{11} & B_{12} \\ 0 & B_{22}\end{bmatrix}\begin{bmatrix}(X^\top\,\Sigma^{-1}\,X)^{-1}\,X^\top\,\Sigma^{-1} \\ (W^\top\,\Sigma^{-1}\,W)^{-1}\,W^\top\,\Sigma^{-1}\end{bmatrix}, \\
&= X\{[B_{11} - I_k\,\theta]\,(X^\top\,\Sigma^{-1}\,X)^{-1}\}\,X^\top\,\Sigma^{-1} \\
&\quad + X\{B_{12}(W^\top\,\Sigma^{-1}\,W)^{-1}\}\,W^\top\,\Sigma^{-1} \\
&\quad + W\{[B_{22} - I_{n-k}\,\theta]\,(W^\top\,\Sigma^{-1}\,W)^{-1}\}\,W^\top\,\Sigma^{-1} + I_n\,\theta, \\
&= XCX^\top\,\Sigma^{-1} + XDW^\top\,\Sigma^{-1} + WEW^\top\,\Sigma^{-1} + I_n\,\theta
\end{aligned} \tag{14}$$

puisque $P_\Sigma = X(X^\top\,\Sigma^{-1}\,X)^{-1}\,X^\top\,\Sigma$ et $(I_n - P_\Sigma) = W(W^\top\,\Sigma^{-1}\,W)^{-1}\,W^\top\,\Sigma^{-1}$ s'additionne à $I_n\,\theta$, C, D et E représentant les matrices entre les accolades correspondantes $\{\cdot\}$. Donc, d'après (14)

$$\Lambda = XCX^\top + XDW^\top + WEW^\top + \Sigma\,\theta$$

pour des matrices appropriées C, D et E ainsi que le scalaire θ choisi pour assurer que $\Lambda \in \mathbf{N}^n$. Ceci requiert essentiellement que D soit nulle, tandis que Λ doit être exprimée comme

$$\Lambda = XCX^\top + WEW^\top + \Sigma\,\theta. \tag{15}$$

Étant donné (15), $P_\Sigma\,\Lambda\,\Sigma^{-1}\,P_\Sigma$ est donné par

$$P_\Sigma\,\Lambda\,\Sigma^{-1}\,P_\Sigma = XCX^\top\,\Sigma^{-1}\,P_\Sigma + 0 + \theta\,P_\Sigma = \Lambda\,\Sigma^{-1}\,P_\Sigma$$

et, d'après le théorème 2, $\mathcal{X}$ est invariant sous $\Lambda\,\Sigma^{-1}$ comme requis [*cf.* Rao, 1967 : 364, équation (6.8)]. L'équation (15) est donc nécessaire et suffisante pour que $\mathcal{X}$ soit invariant sous $\Lambda\,\Sigma^{-1}$, ce qui est nécessaire et suffisant pour que $\widehat{\mu} = \widetilde{\mu} \Leftrightarrow \beta = \widehat{\beta}$; d'où Λ ne joue aucun rôle dans l'estimation efficiente de μ ou de β. Pour exprimer la situation autrement, si Λ suit (15) donc $P_\Sigma = P_\Omega$. Bien entendu E dans (15) peut être nulle et C peut induire l'élimination de quelques colonnes de X. Enfin, θ peut être nul.

Si $\mathcal{X}$ est invariant sous Λ et sous Σ, alors il est invariant sous $\Omega = (\Lambda + \Sigma)$. D'après le théorème 3, $\widetilde{\mu} = \widehat{\mu}$ si et seulement si $\Lambda\,\Sigma^{-1}\,\mathcal{X} \subset \mathcal{X}$, et par le théorème 1, $\widehat{\mu} = \mu^*$ si et seulement si $\Sigma^{-1}\,\mathcal{X} = \mathcal{X}$. D'où $\widetilde{\mu} = \widehat{\mu} = \mu^*$ si et seulement si $\Lambda\,\Sigma^{-1}\,\mathcal{X} \subset \Sigma^{-1}\,\mathcal{X} = \mathcal{X}$.

2. Applications

2.1 *Erreurs équicorrélées*

Un cas simple de régression linéaire qui suit de (1) est lorsque la matrice X contient, en tant qu'une de ses colonnes, le vecteur équiangulaire e dans $\mathbf{R}^n$ [c'est-à-dire, le vecteur $(n \times 1)$ avec tous les éléments égaux à l'unité], $\Sigma = I_n\,\sigma^2(1-\rho)$ et $\Lambda = ee^\top\,\rho\,\sigma^2$, $0 < \sigma^2 < \infty$, $0 < \rho < 1$. Dans ce cas $\Lambda + \Sigma = \Omega$ est une matrice de dimension $(n \times n)$ avec les éléments diagonaux égaux à σ^2 et les éléments non diagonaux égaux à $\rho\,\sigma^2$. Donc, toutes les paires des erreurs ε_i et ε_j $(i, j = 1, 2, ..., n;$ $i \neq j)$ sont identiquement corrélées avec la même variance pour chaque erreur. $\Lambda\,\Sigma^{-1} = ee^\top\,\theta$ où $\theta = \left\{\dfrac{\rho}{1-\rho}\right\}$. Puisque $e \in \mathcal{X}$, $Fe = e$ pour tout $F \in \mathbf{P}$, par conséquent, $Fee^\top\,\theta\,F = ee^\top\,\theta\,F$. D'après le théorème 3, $\mathcal{X}$ est invariant sous $\Lambda\,\Sigma^{-1}$ et $\widehat{\mu} = \widetilde{\mu}$. Par ailleurs, comme $\Sigma = I_n\,\sigma^2(1-\rho)$, il est clair que $\mu^* = \widehat{\mu} = \widetilde{\mu}$. Ce simple exemple est généralement utilisé pour illustrer l'invariance du sous-espace de RCF $\mathcal{X}$ de (1) sous la matrice de dispersion des erreurs Ω. Puisque e est une colonne, disons la première de X, $\mathcal{X}$ est invariant sous Λ; $\mathcal{X}$ est de même invariant sous Σ; par conséquent, $\mathcal{X}$ est invariant sous la somme $\Lambda + \Sigma = \Omega$. En plus, $\mathcal{X}$ est aussi invariant sous $\Lambda\,\Sigma^{-1}$. Mais supposons que Σ ne représente pas $I_n\sigma^2(1-\rho)$, mais représente une matrice diagonale d'éléments positifs et inégaux ou bien une matrice générale définie positive. Dans ce cas, $\Omega = ee^\top\,\lambda + \Sigma$, $\lambda > 0$ et $\widehat{\mu} = X(X^\top\,\Omega^{-1}\,X)^{-1}\,X^\top\,\Omega^{-1}\,y = X(X^\top\,\Sigma^{-1}\,X)^{-1}\,X^\top\,\Sigma^{-1}\,y = \widetilde{\mu}$ puisque pour tout $F \in \mathbf{P}$, $Fee^\top\,\lambda\,\Sigma^{-1}\,F = ee^\top\,\lambda\,\Sigma^{-1}\,F$. Alors, $\mathcal{X}$ est invariant sous $ee^\top\,\lambda\,\Sigma^{-1}$; $\mathcal{X}$ est aussi invariant sous $ee^\top\,\lambda$, mais $\mathcal{X}$ n'est pas invariant sous Σ. Bien entendu, $\Lambda = ee^\top\,\lambda$ $= XCX^\top + WEW^\top + \Sigma\,\theta$ ou E est une matrice $(n \times n)$ nulle, $\theta = 0$ et C est une matrice $(n \times n)$ avec λ dans la première position et des zéros ailleurs. Par conséquent, Λ satisfait l'équation (15) comme requis. En poussant l'analyse, soit n un nombre pair et Σ une matrice diagonale avec des éléments $0 < \sigma_{ii} < \infty$, $i = 1, 2, 3,$ $..., n$. Soit w la première colonne de W tel que $w = [\sigma_{11}, -\sigma_{22}, \sigma_{33}, -\sigma_{44}, ..., -\sigma_{nn}]^\top$. Donc, $e^\top\,\Sigma^{-1}\,w = 1 - 1 + 1 - 1... + 1 - 1 = 0$ indiquant que $w \in \mathcal{W}$, le complément orthogonal de $\mathcal{X}$ par rapport à $(\cdot, \Sigma^{-1}\,\cdot)$. En suivant (15), soit

$$\Lambda = ee^\top\,\lambda + ww^\top\,\alpha + \Sigma\,\theta = XCX^\top + WEW^\top + \Sigma\,\theta$$

pour une certaine valeur $\alpha > 0$, C et E étant des matrices $(n \times n)$ avec des valeurs respectives de λ et α dans les premières positions et des zéros ailleurs; et $\theta > 0$. Donc,

$$\Lambda \Sigma^{-1} = \lambda\, ee^\top \Sigma^{-1} + \alpha\, ww^\top \Sigma^{-1} + I_n\theta$$

et $\mathcal{X}$ se révèle être invariant sous $\Lambda \Sigma^{-1}$ parce que

$$P_\Sigma \Lambda \Sigma^{-1} P_\Sigma = (\lambda\, ee^\top \Sigma^{-1} + \alpha\, ww^\top \Sigma^{-1} + I_n\theta)\, P_\Sigma,$$

par construction $P_\Sigma\, e = e$ et $w^\top \Sigma^{-1} X = 0$. Donc, ce simple exemple d'une régression linéaire avec des erreurs équicorrélées peut être élargi, en utilisant l'équation (15), pour inclure des strucutures de dispersion plus complexes de Λ et Σ tel que $\mathcal{X}$ est invariant sous $\Lambda \Sigma^{-1}$ et $\hat{\mu}$ et $\tilde{\mu}$ sont égaux.

2.2 *Régression linéaire à coefficients stochastiques*

Une deuxième application du théorème 2 est déduite du modèle RCA tel que décrit par Pfefferman (1984). Ce modèle utilise (1) et (2) mais ajoute quelques conditions supplémentaires, donnant le modèle de (16)-(19) ci-dessous.

$$y = X\beta + \varepsilon, \tag{16}$$

$$\beta \sim [M\gamma, \Delta], \quad \Delta \in \mathbf{Q}^k, \tag{17}$$

$$\varepsilon \sim [0, \Sigma], \quad \Sigma \in \mathbf{Q}^n, \tag{18}$$

$$E[\varepsilon\, \beta^\top] = 0. \tag{19}$$

Dans les expressions (16) et (17), y et ε sont de nouveau des vecteurs aléatoires qui varient dans $\mathbf{R}^n$, ε suivant la distribution (18); en (17), β varie dans $\mathbf{R}^k$ avec une moyenne $E[\beta] = M\gamma$ où M est une matrice fixe et connue de dimension $(k \times p)$ et de rang p et γ est un vecteur de coefficient de $\mathbf{R}^p$ fixe et inconnu. Il s'ensuit que β peut être exprimé comme

$$\beta = M\gamma + \nu, \tag{20}$$

$$\nu \sim (0, \Delta), \quad E[\varepsilon\, \nu^\top] = 0. \tag{21}$$

La littérature portant sur l'estimation de ce modèle s'intéresse généralement aux combinaisons linéaires de β et γ. Plusieurs cas sont revus dans Pfefferman (1984) et étendus, dans un environnement libre de coordonnées, dans Fisher et Voia (2002). Dans le cas d'une estimation efficiente de γ, (20) est remplacée par (16) en tant que prélude pour obtenir les estimateurs MCG :

$$y = XM\gamma + (X\nu + \varepsilon), \tag{22}$$

$$(X\nu + \varepsilon) \sim [0, X\Delta X^\top + \Sigma], X\Delta X^\top \in \mathbf{N}^n. \tag{23}$$

Selon la notation de la première section, $X\Delta X^\top = \Lambda$ de rang k. Suivant (2), le sous-espace de régression de (22) est

$$\mathcal{X}_0 = \{x \in \mathbf{R}^n : x = XM\lambda, \lambda \in \mathbf{R}^p\}$$

et bien entendu $\mathcal{X}_0 \subseteq \mathcal{X}$ avec dim $\mathcal{X}_0 = p \leq k$. Lorsque $p = k$, M est non singulière et $\mathcal{X}_0$ coïncide avec $\mathcal{X}$.

Pour $\Omega = X \Delta X^\top + \Sigma$, l'estimateur MCG de $XM\gamma = \mu$ est,

$$\widehat{\mu} = XM[M^\top X^\top \Omega^{-1} XM]^{-1} M^\top X^\top \Omega^{-1} y = P_{0\Omega} y \tag{24}$$

où $P_{0\Omega} = XM[M^\top X^\top \Omega^{-1} XM]^{-1} M^\top X^\top \Omega^{-1}$ est la matrice de projection sur $\mathcal{X}_0$ orthogonale au produit scalaire $(\cdot, \Omega^{-1} \cdot)$. Soit $p = k$. Donc $\mathcal{X}_0 = \mathcal{X}$ et $P_\Omega = P_{0\Omega}$. De plus, selon le théorème 3, $\mathcal{X}$ est invariant sous $\Lambda\, \Sigma^{-1}$. D'où $P_{0\Omega} = P_\Omega = P_\Sigma$ et

$$\widetilde{\mu} = X(X^\top \Sigma^{-1} X)^{-1} X^\top \Sigma^{-1} y = P_\Sigma y = P_\Omega y = \widehat{\mu}.$$

Selon $\widehat{\mu}$ dans (24),

$$\begin{aligned}\widehat{\gamma} &= (M^\top X^\top \Omega^{-1} XM)^{-1} M^\top X^\top \Omega^{-1} y \\ &= M^{-1}(X^\top \Sigma^{-1} X)^{-1} X^\top \Sigma^{-1} y\end{aligned}$$

d'où $\widehat{\gamma} = M^{-1} \widehat{\beta}$, ce qui implique que $\widehat{\beta} = M\widehat{\gamma}$ et que $\widehat{\beta}$ et $\widehat{\gamma}$ sont indépendants de Δ. Rao (1965) a étudié le cas où $M = I_k$ ce qui donne $\widehat{\beta} = \widehat{\gamma}$ les deux étant indépendants de Δ; si, en plus $\Sigma = I_n\, \sigma^2$, les estimateurs MCO de β et γ sont efficients et indépendants de Δ et de σ^2. Plus généralement, lorsque M est non singulière et $\mathcal{X}$ est invariant sous Σ, alors $\mu^* = \widehat{\mu} = \widetilde{\mu} = Py$, l'estimateur MCO de μ; les estimations efficientes de μ, β et γ sont tous indépendants de Δ et Σ.

Si M est rectangulaire, alors XM couvre $\mathcal{X}_0$, qui est un sous-espace approprié de $\mathcal{X}$, donc le sous-espace de régression possède une dimension inférieure au rang de $X \Delta X^\top$. Une telle condition assure que $\mathcal{X}_0$ n'est pas invariant sous $X \Delta X^\top$ et par conséquent n'est pas invariant sous $X \Delta X^\top \Sigma^{-1}$. Dans ce cas, l'estimation efficiente de γ dépend certainement de Δ et de Σ.

2.3 *Un modèle linéaire mixte*

Un troisième exemple illustratif requiert la répartition de X dans (16) en $[X_1 : X_2]$ où X_1 contient k_1 colonnes et X_2 contient k_2 colonnes ($k_1 + k_2 = k$). Le vecteur de coefficient β est réparti de même en β_1 et β_2. Les expressions (16) et (17) sont donc remplacées par

$$y = X_1 \beta_1 + X_2 \beta_2 + \varepsilon, \tag{25}$$

$$\beta_2 \sim [M_2 \gamma_2, \Delta_2], \quad \Delta_2 \in \mathbf{Q}^{k_2} \tag{26}$$

et β_1 est considéré comme un vecteur de coefficient fixe mais inconnu. Le sous-espace couvert par X_1 est $\mathcal{X}_1$ et celui de X_2 est $\mathcal{X}_2$ et, par l'indépendance linéaire de X_1 et X_2, $\mathcal{X}_1 \oplus \mathcal{X}_2 = \mathcal{X}$. La matrice M_2 est de dimension $(k_2 \times p_2)$ et de rang $p_2 \leq k_2$. Les équations (20) et (21) deviennent

$$\beta_2 = M_2 \gamma_2 + \nu_2 \tag{27}$$

$$\nu_2 \sim [0, \Delta_2], \quad E[\varepsilon\, \nu_2^\top] = 0 \tag{28}$$

et (22) devient

$$y = X_1 \beta_1 + X_2 M_2 \gamma_2 + (X_2 \nu_2 + \varepsilon),$$

$$(X_2 \nu_2 + \varepsilon) \sim [0, X_2 \Delta_2 X_2^\top + \Sigma].$$

Pour ce modèle mixte à coefficients de régression fixes et stochastiques, les résultats peuvent être résumés comme suit :

1. Si M_2 est non singulière, $p_2 = k_2$ et l'estimation efficiente de β_2 et γ_1 est indépendante de Δ_2 : si $\mu = X_1 \beta_1 + X_2 M_2 \gamma_2$, donc l'estimateur efficient est $\widehat{\mu} = X(X^\top \Sigma^{-1} X)^{-1} X^\top \Sigma^{-1} y = Py$.
2. Si en plus de la non-singularité de M_2, $\mathcal{X}$ est invariant sous Σ, alors $\mu^* = \widehat{\mu} = X(X^\top X)^{-1} X^\top y$, et l'estimation est indépendante de Δ_2 et de Σ.
3. Si $p_2 < k_2$, l'estimation efficiente de β_1 et de γ_2 dépend de Δ_2 et Σ.

Ces résultats peuvent être obtenus en notant que

$$X_2 \Delta_2 X_2^\top = [X_1 \quad X_2] \begin{bmatrix} 0 & 0 \\ 0 & \Delta_2 \end{bmatrix} \begin{bmatrix} X_1^\top \\ X_2^\top \end{bmatrix}.$$

La dernière expression peut être écrite sous la forme $XCX^\top = \Lambda$, C étant la matrice centrale à droite de l'égalité ci-dessus. En plus,

$$[X_1 \quad X_2 M_2] = X \begin{bmatrix} I & 0 \\ 0 & M_2 \end{bmatrix} = XM$$

où M est une matrice diagonale par blocs de blocs I et M_2. Ainsi, les résultats de cette section sont essentiellement identiques à ceux de la section 2.2, car $XM\gamma = X_1 \beta_1 + X_2 M_2 \gamma_2$, $\gamma = [\beta_1^\top \quad \gamma_2^\top]^\top$, et $(X_2 \nu_2 + \varepsilon) \sim [0, XBX^\top + \Sigma]$. Il s'ensuit que la condition de McAleer (1992), stipulant que l'estimation efficiente requiert l'orthogonalité de X_1 et X_2, n'est pas juste.

3. L'estimation directe et à deux étapes

En se reférant au modèle (16)-(19), lorsque $M = I_k$ et donc $E[\beta] = \gamma$, Rao (1965) montre que l'estimateur sans biais à variance minimale de γ est

$$\tilde{\gamma} = (X^\top \Sigma^{-1} X)^{-1} X^\top \Sigma^{-1} y.$$

Dans ce cas, $\widehat{\beta}$ qui est obtenu par l'estimation MCG directe de (16), est identique à $\tilde{\gamma}$ qui est indépendant de Δ. Ce résultat implique que $X\tilde{\gamma} = P_\Sigma y = X\widehat{\gamma} = P_\Omega y$, car $\tilde{\gamma} = (X^\top \Omega^{-1} X)^{-1} X^\top \Omega^{-1} y$ où $\Omega = (X \Delta X^\top + \Sigma)$ est l'estimateur MCG naturel de γ dans l'equation consolidée (22) et, par conséquent, doit donner lieu à l'estimateur linéaire sans biais à variance minimale; toutefois, Rao (1965) n'offre pas de démonstration directe de l'égalité entre $\widehat{\gamma}$ et $\tilde{\gamma}$.

Lorsque M n'est pas nécessairement égale à I_k, mais peut être n'importe quelle matrice non singulière connue d'ordre k, alors $\mathcal{R}[XM] = \mathcal{X}$ et puisque $\mathcal{X}$ est invariant sous $X\,\Delta\,X^\top\,\Sigma^{-1}$,

$$XM\widehat{\gamma} = P_{0\Omega}y = P_\Omega y = P_\Sigma y = XM\widetilde{\gamma} = X\widehat{\beta}.$$

Donc, l'estimateur linéaire sans biais à variance minimale de $E[X\,\beta]$ est $XM\widehat{\gamma} = XM\widetilde{\gamma} = X\widehat{\beta}$, où tous ces termes sont indépendants de Δ.

Lorsque M est une matrice connue de dimension $k \times p$ et de rang p, il existe deux méthodes pour estimer γ : directement ou indirectement. Dans la méthode directe l'estimation MCG est appliquée à (22) donnant

$$XM\widehat{\gamma} = P_{0\Omega}y = XM(M^\top X^\top \Omega^{-1} XM)^{-1} M^\top X^\top \Omega^{-1} y. \tag{29}$$

La méthode indirecte contient deux étapes : premièrement β est estimée en appliquant la méthode MCG à (16), ce qui donne

$$X\widehat{\beta} = P_\Sigma y.$$

Ensuite l'équation (20) est reformulée comme

$$\widehat{\beta} = M\gamma + (\widehat{\beta} - \beta) + \nu \tag{30}$$

$$(\widehat{\beta} - \beta) + \nu \sim [0, (X^\top \Sigma^{-1} X)^{-1} + \Delta], \tag{31}$$

et dans la deuxième étape, γ est estimée en appliquant la méthode MCG à (30) tout en utilisant la dispersion donnée en (31). Écrivons $\Gamma = [(X^\top \Sigma^{-1} X)^{-1} + \Delta]$, Γ^{-1} peut être obtenu de l'identité $X^\top = \Gamma^{-1}\,\Gamma\,X^\top = \Gamma^{-1}\,(X^\top \Sigma^{-1} X)^{-1} X^\top \Sigma^{-1}\,\Omega$ et en postmultipliant $X^\top$ et la dernière expression par $\Omega^{-1} X$, $X^\top \Omega^{-1} X = \Gamma^{-1}$. Donc l'estimateur en deux étapes de γ est $\ddot{\gamma}$ qui est défini en

$$XM\ddot{\gamma} = XM(M^\top \Gamma^{-1} M)^{-1} M^\top \Gamma^{-1} \widehat{\beta}$$

où la substitution de Γ^{-1} donne $XM\ddot{\gamma} = P_{0\Omega} X\widehat{\beta} = P_{0\Omega} P_\Sigma y$. Puisque $\mathcal{X}$ est invariant sous $X\,\Delta\,X^\top\,\Sigma^{-1}$, $P_\Sigma = P_\Omega$ et puisque $\mathcal{X}_0 \subseteq \mathcal{X}$, $P_{0\Omega} P_\Omega = P_{0\Omega}$. Donc,

$$XM\ddot{\gamma} = P_{0\Omega} P_\Sigma y = P_{0\Omega} P_\Omega y = P_{0\Omega} y = XM\widehat{\gamma} \tag{32}$$

et tout comme dans XM, $\ddot{\gamma} = \widehat{\gamma}$. Quand $p < k$, l'estimation efficiente sans biais dépend de Δ. Quand $p = k$, M est non singulière et $M\ddot{\gamma} = M\widehat{\gamma} = M\widetilde{\gamma} = \widehat{\beta}$.

Géométriquement, $XM\ddot{\gamma}$ est obtenu, premièrement, en appliquant la projection orthogonale à $(\cdot, \Sigma^{-1} \cdot)$ de y sur $\mathcal{X}$, qui, par invariance, est identique à la projection orthogonale à $(\cdot, \Omega^{-1} \cdot)$: $P_\Sigma y = P_\Omega y$; deuxièmement, en projetant le résultat sur $\mathcal{X}_0$ par rapport au produit scalaire $(\cdot, \Omega^{-1} \cdot)$: $P_{0\Omega} P_\Sigma y = P_{0\Omega} P_\Omega y$. Puisque $\mathcal{X}_0 \subseteq \mathcal{X}$, $P_{0\Omega} P_\Omega y = P_{0\Omega} y$. Les deux méthodes donnent le même résultat pour trois raisons :

(a) $\mathcal{X}$ est invariant sous $X\,\Delta\,X^\top\,\Sigma^{-1}$ et donc $P_\Sigma = P_\Omega$;

(b) les trois projections $P_{0\Omega}$, P_Ω et P_Σ sont toujours dans la même direction, c'est-à-dire orthogonales par rapport à $(\cdot, \Omega^{-1} \cdot)$; et

(c) $\mathcal{X}_0 \subseteq \mathcal{X}$.

CONCLUSION

L'objectif de cet article a été de développer, d'un point de vue différent, un ensemble de résultats pour le modèle général de régression linéaire avec des coefficients stochastiques. Ce faisant les résultats existants ont été revus, précisés et étendus. Les conclusions principales sont les suivantes :

1. Si la dispersion des erreurs dans une régression linéaire peut être exprimée comme $(\Lambda + \Sigma)$, où Λ est définie non négative et Σ est définie positive, tel que la somme est définie positive, une condition nécessaire et suffisante pour que les estimateurs MCG des coefficients soient indépendants de Λ est que $\Lambda\, \Sigma^{-1}\, \mathcal{X} \subseteq \mathcal{X}$ (Théorème 2).

2. Un moyen simple de détecter l'invariance a été développé (Théorème 3 *et seq.*) et illustré par trois exemples.

3. Le modèle général de régression linéaire à coefficients stochastiques a été décrit dans les expressions (16)-(19). Trois résultats principaux ont été développés pour ce modèle :

   (a) Lorsque la matrice M est carrée et non singulière, la variation engendrée par les coefficients stochastiques, $(X\, \Delta\, X^{\top})$, ne joue aucun rôle dans l'estimation par les MCG de la moyenne de y $(XM\,\gamma)$, quelque soit Δ.

   (b) Lorsque (3a) ci-haut s'applique, $\widehat{\beta} = M\,\widehat{\gamma}$ indépendamment de Δ. Rao (1965) a étudié le cas où $M = I_k$ et donc où $\widehat{\beta} = \widehat{\gamma}$, indépendamment de Δ.

   (c) Si en plus, $\mathcal{X}$ est invariant par rapport à la dispersion des erreurs Σ alors les estimateurs des MCG sont equivalents aux estimateurs des MCO et sont, par conséquent, indépendants à la fois de Σ et Δ.

4. Une explication géométrique de l'équivalence entre certains estimateurs a été présentée. Ceci concerne l'équivalence entre l'estimation des MCG directe des coefficients fixes (γ) de (22) et l'estimation des MCG en deux étapes : premièrement, l'estimation des MCG de β dans l'équation (20), deuxièmement, l'estimation des MCG de γ dans l'équation (30) en utilisant l'estimateur de β. En dépit de certaines différences apparentes dans les formules de la dispersion, nous avons démontré que l'estimation directe coïncide avec la projection directe sur $\mathcal{X}_0$ par rapport à $(\cdot, \Omega^{-1} \cdot)$ donnant lieu à $P_{0\Omega}\, y$, tandis que l'estimation en deux étapes implique une projection correspondante sur $\mathcal{X}$ pour donner $P_\Omega\, y = P_\Sigma\, y$, suivi par une projection de $P_\Sigma\, y$ ou $P_\Omega\, y$ sur $\mathcal{X}_0$. Lorsque M est non singulière $P_{0\Omega} = P_\Omega$ et $M\,\widehat{\gamma} = \widehat{\beta}$; lorsque $M = I_k$, $\widehat{\gamma} = \widehat{\beta}$.

BIBLIOGRAPHIE

ANDERSON, T. W. (1948), « On the Theory of Testing Serial Correlation », *Skandinavisk Aktuarietidskift*, 31 : 88-116.

BAKSALARY, J. K. (1988), « Criteria for the Equality between Ordinary Least Squares and Best Linear Unbiased Estimators under Certain Linear Models », *The Canadian Journal of Statistics*, 16 : 97-102.

BAKSALARY, J. K. et G. TRENKLER (1989), « Solution 88.3.4: The Efficiency of OLS in a Seemingly Unrelated Regressions Model », *Econometric Theory*, 5 : 463-465.

BALESTRA, P. (1970), « On the Efficiency of Ordinary Least Squares in Regression Models », *Journal of American Statistical Association*, 65 : 1 330-1 337.

DWIVEDI, T. D et V. K. SRIVASTAVA (1978), « Optimality of Least Squares in the Seemingly Unrelated Regression Equation Model », *Journal of Econometrics*, 7 : 391-395.

FISHER, G. et MARCEL CHRISTIAN VOIA (2002), « Estimating Systems of Stochastic Coefficients Regressions when Some of the Observations are Missing », *in* A. ULLAH, A. T. KWAN et A. CHATURVEDI (éds), *Handbook of Applied Econometrics and Statistical Inference*, chap. 23, New York : Marcel Dekker.

GOURIÉROUX, C. et A. MONFORT (1980), « Sufficient Linear Structures: Econometric Applications », *Econometrica*, 48 : 1 083-1 097.

HALMOS, P. R. (1942), *Finite-Dimensional Vector Spaces*, Édition préliminaire, Princeton : Princeton University Press.

HALMOS, P. R (1958), *Finite Dimensional Vector Spaces*, 2e édition, Princeton : van Nostrand.

KAPTEYN, A et D. G. FIEBIG (1981), « When are Two-stage and Three-stage Least Squares Estimators Identical? », *Economics Letters*, 8 : 53-7.

KLOEK, T. (1981), « OLS Estimation in a Model Where a Microvariable is Explained by Aggregates and Contemporaneous Disturbances are Equicorrelated », *Econometrica*, 49 : 205-207.

KRUSKAL, W. (1968), « When are Gauss-Markov and Least Squares Estimators Identical? A Coordinate-free Approach », *Annals of Mathematical Statistics*, 39 : 70-75.

LONGFORD, N. T. (1993), *Random Coefficients Models, Oxford* : O. U. P.

MCALEER, M. (1992), « Efficient Estimation: The Rao-Zyskind Condition, Kruskal's Theorem and Ordinary Least Squares », *Economic Record*, 68 : 65-72.

PFEFFERMAN, D. (1984), « On Extensions of the Gauss-Markov Theorem to the Case of Stochastic Regression Coefficients », *Journal of the Royal Statistical Society*, B, 46 : 139-148.

PHILOCHE, J. -L. (1971), « À propos du théorème de Gauss-Markov », *Annales de l'Institut Henri Poincaré*, 7 : 271-281.

PUNTANEN, S. et G. P. H. STYAN (1989), « The Equality of the Ordinary Least Squares Estimator and the Best Linear Unbiased Estimator », *The American Statistician*, 43 : 153-161.

RAO, C. R. (1965), « The Theory of Least Squares when the Parameters Are Stochastic and its Application to the Analysis of Growth Curves », *Biometrika*, 52 : 447-458.

RAO, C. R. (1967), « Least Squares Theory Using an Estimated Dispersion Matrix and its Application to Measurement of Signals », ***in*** *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Volume 1, Berkeley : U. C. P., p. 355-372.

SRIVASTAVA, V. K. et D. E. A. GILES (1987), *Seemingly Unrelated Regression Equations Models*, New York : Marcel Dekker.

SWAMY, P. A. V. B. (1971), *Statistical Inference in Random Coefficient Regression Models*, Berlin : Springer-Verlag.

SWAMY, P. A. V. B., R. K. Conway et M. R. LeBlanc (1988a), « The Stochastic Coefficients Approach to Econometric Modeling, part I: A Critique of Fixed Coefficients Models », *The Journal of Agricultural Economic Research*, 40 : 2-10.

SWAMY, P. A. V. B., R. K. CONWAY et M. R. LEBLANC (1988b), « The Stochastic Coefficients Approach to Econometric Modeling, part II: Description and Motivation », *The Journal of Agricultural Economic Research*, 40 : 21-30.

SWAMY, P. A. V. B., R. K. Conway et M. R. LeBlanc (1989), « The Stochastic Coefficients Approach to Econometric Modeling, part III: Estimation, Stability Testing and Prediction », *The Journal of Agricultural Economic Research*, 41 : 4-20.

WATSON, G. S. (1967), « Linear Least Squares Regression », *The Annals of Mathematical Statistics*, 38 : 1 679-1 699.

WATSON, G. S. (1972), « Prediction and the Efficiency of Least Squares », *Biometrika*, 59 : 91-98.

ZYSKIND, G. (1962), « On Conditions for Equality of Best and Simple Least Squares Estimators », abstract, *The Annals of Mathematical Statistics*, 33 : 1 502-1 503.

ZYSKIND, G. (1967), « On Canonical Forms, Non-negative Covariance Matrices and Best and Simple Least Squares Linear Estimators in Linear Models », *The Annals of Mathematical Statistics*, 38 : 1 092-1 109.